Vol. 205, No. 1, 1994 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
November 30, 1994 Pages 490-496

Structure of the Human Hexokinase II Gene*

Mari Malkki, Markku Laakso and Samir S. Deeb†

Departments of Medicine and Genetics, University of Washington, Seattle, Washington 98195

Received August 22, 1994

SUMMARY Mutations in the gene encoding hexokinase II which catalyzes a key step in glycolysis could contribute to the development of peripheral insulin resistance and lead to non-insulin-dependent diabetes mellitus. As a first step towards screening patients for mutations in this gene, we have determined its structure and the sequence of exon-intron junctions. The human HKII gene is composed of 18 exons that span at least 40 kb, and its organization is highly homologous to that of the rat gene. A hexokinase II processed pseudogene was discovered while screening a human genomic library. The coding sequence of this pseudogene is uninterrupted by introns and contains at least one premature stop codon.

Hexokinase (ATP:D-hexose 6-phosphotransferase, EC 2.7.1.1) isozymes I-IV catalyze the first rate-limiting step in glycolysis, converting glucose to glucose-6-phosphate. Isoenzymes I-III in mammalian tissues consist of a single polypeptide chain of 100 kDa, have high affinity for glucose and are subject to allosteric inhibition by glucose-6-phosphate. In contrast, type IV hexokinase (50 kDa), commonly referred to as glucokinase (GK), is produced by mammalian liver and pancreas, has relatively low affinity for glucose, and is not subject to product inhibition (reviewed in 1).

Hexokinase II (HKII) is present in insulin responsive tissues such as adipose and skeletal muscle (2,3). Sequences of the human and the rat HKII have been deduced from the nucleotide sequence of cDNA clones isolated from human and rat skeletal muscle and adipose tissue libraries (4, 5, 6). Human HKII has 94 % identity with rat HKII and 72% identity with human HKI (6). The N- and C-terminal halves of the hexokinases I-III show extensive sequence similarity (4,6-

* This work was supported by U.S. Public Health Service Grant HL-30086.

† Corresponding author. FAX: 206 543 0754.

Abbreviations: Hexokinase, HK; glukokinase, GK; non-insulin-dependent diabetes mellitus, NIDDM.

0006-291X/94 $5.00

9) to the 50 kDa GK and the yeast hexokinase, supporting the view that hexokinases I-III had evolved from a hexokinase ancestral gene by gene duplication followed by tandem fusion (10,11). The rat HKII gene is composed of 18 exons and spans at least 41 kb of DNA (5). The human HKII gene has been localized to chromosome 2p13.1(12).

Insulin induces rat HKII mRNA and activity levels in adipose and skeletal muscle (5,13). Non-insulin-dependent diabetes mellitus (NIDDM) is characterized by reduced rates of both glucose oxidation and nonoxidation in adipose and muscle tissues and by relatively low levels of glucose-6-phosphate (14,15,16). Mutations in the gene encoding glucokinase were shown to lead to the development of NIDDM (reviewed in 17). Therefore, we have begun to assess the role of mutations in the HKII gene in NIDDM. To this end, we have determined the intron-exon structure of the human HKII gene in order to facilitate screening it for mutations among individuals with insulin resistance and NIDDM.

MATERIALS AND METHODS

Amplification of introns. PCR amplification of DNA fragments containing introns was conducted in a 100 ul volume using the Perkin Elmer Cetus GeneAMP kit with 100 ng genomic DNA and 50 pmol of each primer. For some reactions (introns 4, 8-11, 13, 14) 10% glycerol was included. PCR conditions were: denaturation at 94 ^{o}C for 2 min followed by 35 cycles of denaturation at 92-94^{o}C for 0.5 -1 min, annealing at 60-66o C for 45-60 sec and extension at 72^{o}C for 0.5 - 2.5 min with a final extension at 72^{o}C for 4 min. Amplified fragments were purified by electrophoresis on a low-melting point agarose and directly used as templates to determine the sequence of exon-intron junctions (Sequenase kit, US Biochemicals, Cleveland, OH) as previously described (18).

Isolation of genomic clones. Genomic clones were isolated from a human lung fibroblast genomic library prepared in bacteriophage λ FIX ^{TM} (Stratagene #944201, LaJolla, CA). *E.coli* XLI-Blue MRA (P2) (Stratagene, #200303) was used as host. Three segments from the 5' region of human HKII cDNA were used to screen the library. Probe A: 286 bp fragment (-52 to +234 with respect to the translation start site); probe B: 112 bp fragment (-52 to +60); probe C: 187 bp fragment (+ 296- + 483). Approximately 25,000-30,000 plaques were plated per 150 mm LB broth plate and incubated for 10 h at 37^{o}C. Plaque lifts (Hybond N membrane, Amersham) were prepared according to the manufacturer's protocol. Membranes were hybridized to ^{32}P-labeled (random primed labeling kit, Boehringer Mannheim, specific activity approximately 1x10^{9} cpm/μg) human HK II cDNA probe A, B or C at 42^{o}C overnight. Membranes were washed 2 times in 0.5xSSC-0.5% SDS at 65^{o}C for 30 minutes and autoradiographed at -70^{o}C overnight. After three rounds of plating and hybridization, positive plaques were used to make stocks and phage DNA from liquid lysates as described (19). The phage DNA was used directly as template in PCR amplification reactions. Cycle sequencing was performed using the "fmol DNA sequencing system" (Promega, Madison, WI) according to the manufacturer's protocol.

RESULTS

Determination of exon-intron junctions by PCR amplification across introns

Guided by the exon-intron structure of the rat HKII gene (5), oligonucleotide primers complementary to human HKII coding sequences (6) were used to amplify across introns using total human genomic DNA as template. Amplified DNA segments were directly used as templates to determine the sequence of exon-intron junctions using the exonic oligonucleotides as primers. The results are shown in Table 1. The size and position of exons together with the codons that are interrupted by introns are given in Table 2. Introns interrupted coding sequences at the same nucleotide positions as those observed in the rat HKII gene (5). The length of introns 4 , 5, 7-11, and 15-17 (Figure 1) was determined by gel electrophoresis of amplified fragments that contained an intron flanked by coding sequences. The length of introns 6 (322 bp), 11(259 bp), and 14 (96 bp) was determined by sequencing. The structural organization of exons 1-4 could not be determined using the above strategy since introns 1-3 could not be amplified using total genomic DNA as template. Therefore, genomic clones containing these sequences were isolated.

Characterization of genomic clones containing exons 1 - 4 and a pseudogene

A total 2.5 million recombinant phage particles of a human genomic library were screened with three human HKII cDNA probes A, B , and C which are derived

Table 1. Sequence of the exon-intron junctions of the human HK II gene +

Number	5'- Splice donor		3'-splice acceptor	
1	AAG	gtaagtcagcgcgggcggggcggca.....	.....tgctcttgtcttcctccttttttcag	GTT
2	CAG	gtactgcatctgggggatggctcta.....	.....atagtggcccttcctttctctgcag	AAC
3	CAG	gtatgacccttctctcagggcagcc.....	.....accaccctgtgtcatgtatccttag	CTG
4	GAG	gtaagatgggctcctcagacacttg.....	.....tattcacttcttggtccctttccag	AGT
5	GGG	gtgagtggggtggcaggagcttggg.....	.....agagcctcccatttgtctccacag	GAC
6	TGG	gtgagtgaacaccgtgcatgaaggg.....	.....gcatagccgtcccttgttttggcag	GCA
7	ACT	gtgagtaggcccttcctgtgcgagt.....	.....gtttttggtttgtgtcttccggcag	GTT
8	AGG	gtgagcttctggccagcccccctcta.....	.....cacatgctgtctttctgtttcccag	GGA
9	CCA	gtgagtcagtgtgcagggcctggag.....	.....tcttacccgccctggggaactgcag	TTT
10	CAG	gtacacggcagggttgccacctggc.....	.....tcacttccctgggcttatttttccag	AGA
11	GAG	gtgagcagggcgcgccttcaggagg.....	.....tcagtgtcctaacttctccctgcag	CTC
12	GAG	gtaacagcaccttcctggagggctc.....	.....ttccatgcttgtgtgtgatttttag	AGC
13	GAG	gtaggagacacatggcacgaggtct.....	.....tgggcatctgcttcttccctctcag	GAG
14	TTG	gtaaggacccagactgtcctttcca.....	.....ggccctggtatcctgtgattggcag	GCA
15	GAG	gtaggcacccaactggggccctgtt.....	.....tccatgaatgttacttgcattgcag	GTT
16	GAG	gtgagagcttagggctcagggtagc.....	.....acttgctgtctccctcccaccccag	TGA
17	TCA	gtgagtgcctgatcccagccctccc.....	.....tgtgtcttcctcccaccttttccag	CTT

+ Exon sequences in upper case and intron sequences in lower case.

Table 2. Position and size of exons in the human HK II gene †

Number	Position and length (bp)¶	Codons interrupted	Amino acid and position
1	1-63 (63)	AAG I GTT	Lys^{21}-Val^{22}
2	64-226 (163)	G I AA	Glu^{76}
3	227-375 (149)	CAG I CTG	Gln^{135}-Leu^{136}
4	376-495 (120)	GAG I AGT	Glu^{165}-Ser^{166}
5	496-591 (96)	GGG I GAC	Gly^{197}-Asp^{198}
6	592-691 (100)	G I GC	Gly^{231}
7	692-875 (184)	CT I G	Leu^{292}
8	876-1031 (156)	GG I G	Gly^{344}
9	1032-1265 (234)	CA I T	His^{422}
10	1266-1570 (305)	G I AG	Glu^{524}
11	1571-1719 (149)	GAG I CTC	Glu^{573}-Leu^{574}
12	1720-1839 (120)	GAG I AGC	Glu^{613}-Ser^{614}
13	1840-1935 (96)	GAG I GAG	Glu^{645}-Glu^{646}
14	1936-2035 (100)	G I GC	Gly^{679}
15	2036-2219 (184)	AG I G	Arg^{740}
16	2220-2375 (156)	AG I T	Ser^{792}
17	2376-2609 (234)	CA I C	His^{870}
18	2610-2751 (142)		

† Positions of the exons and amino acid residues are relative to the initiation codon (6). ¶The length of coding sequences only is given for exons 1 and 18. Vertical bars indicate the positions of introns.

from the 5' region of HKII cDNA (see Materials and Methods). A total of 24 clones were isolated, 8 of which using probe A, 9 using probe B, and 7 using probe C.

Six (two were identical) of the 8 clones that hybridized to probe A hybridized to probe B (the 5' segment of probe A) as well. DNA from four of the six clones that hybridized to probes A and B was digested with either SstI or XbaI and subjected to Southern blot analysis. Restriction fragments that hybridized to probe A were subcloned into the pGEM 3Zf (+) vector (Promega, Madison, WI) and used to determine the sequence of exon-intron junctions using the following primers that are derived from the 5' coding region of human HKII cDNA : 5'CTCTGCGTCTCCGCCTCGG3', 5'CTCACCCTACTGCAGCAGTG3', 5'CCATCTGGAGTGGACCTCAC3' (6). Surprisingly, all of these subclones contained a sequence that was very homologous to that of human HKII cDNA (6) but was uninterrupted by intronic sequences. Moreover, the cding sequence in these clones contained at least one premature stop codon (at residue 120). Therefore, these clones were most likely derived from the same processed HKII pseudogene.

The two clones that hybridized only with probe A were identical as determined by restriction enzyme analysis. An EcoRI fragment of one of these clones (HKII

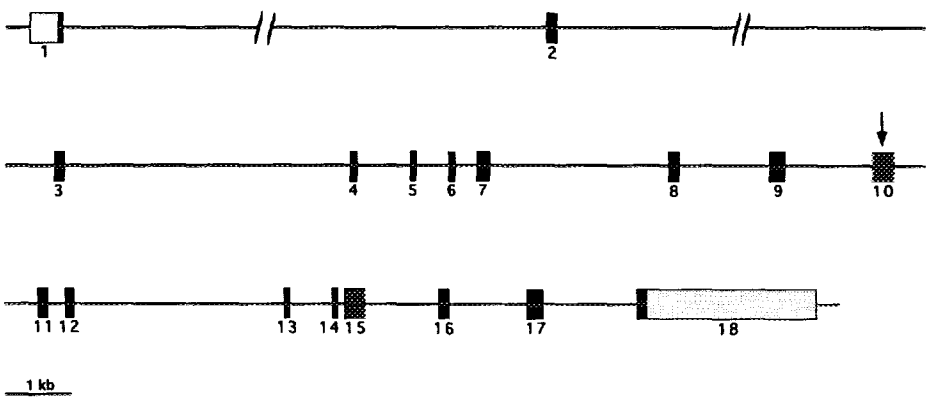

Figure 1. Structure of the human HKII gene. The 18 exons are represented by filled boxes and introns by a thin line. Coding sequences are represented by black boxes and noncoding sequences by dotted boxes. The arrow indicates the proposed point of fusion of two ancestral HK genes to form the HKII gene (5). The breaks in introns 1 and 2 indicate that the length of these introns is unknown. The transcription initiation site has not been determined and therefore the exact length of the 5'-untranslated region is unknown.

L3), which was shown by Southern analysis to contain coding sequences, was subcloned and used as template to determine the sequence of exon-intron junctions. The results showed that this fragment contained exon 2 and the flanking intron sequences. The junctions occurred at the same nucleotide positions as those observed in the rat HKII gene (5). No other coding sequences were detected in the original phage clone which contained an insert of approximately 20 kb.

In order to isolate exon 1-containing clones, the phage genomic library was next screened using probe B alone. Six clones were isolated and examined for the presence of HKII sequences using DNA from phage lysates as template to amplify and sequence a segment (-101 to + 60) that flanked the translation initiation site. Three of these clones were shown to contain the same pseudogene sequences described above. However, three contained sequences identical to those in the 5' coding region of HKII cDNA. DNA from the clones that contained authentic HKII gene sequences was used as template to determine the sequence of the exon 1-intron 1 junction as described under Methods. This junction occurred at the same nucleotide position as that observed in the rat HKII gene (5).

The organization of exons 3 and 4 was determined using genomic clones that hybridized to probe C, which contained sequences from exons 3 and 4 (according to rat HKII). DNA from plate lysates of seven such clones was used as template to amplify a segment containing intron 3 and the flanking exon 3 and 4 sequences. Using a pair of primers from exon 3 and 4 , a single PCR-

fragment (4.9 kb) was obtained with clone L11 DNA as template. This fragment was used to deduce the length of intron 3 (4.47 kb) and to determine the sequence of the exon 3-intron 3 and intron 3 exon-4 junctions. Using the same clone as template with the T3 (vector-specific primer in close proximity with the cloning site) and the exon 3 primer 5' CTCAGGGATGGCATAGATCTG 3', a segment (3.0 kb) containing the intron 2-exon 3 junction was amplified and used to determine the sequence flanking this junction. The results showed that exons 3 and 4 are identical in size to the corresponding exons of the rat HKII gene. A Diagram of the intron-exon structure of the human HKII gene is shown in Figure 1.

DISCUSSION

The structural organization of the human HKII gene has been determined by analysis of genomic clones and by PCR amplification across introns. The gene, which spans at least 40 kb, is comprised of 18 exons that range in size from 63 bp to 305 bp. The length of introns 3-17 ranged between 96 bp and 4.5 kb. However, the length of introns 1 and 2 could not be determined by either PCR amplification or examination of genomic clones. Therefore, as in the rat HKII gene, introns 1 and 2 are likely to be more than 10 kb in length. The structure of the human HKII is highly homologous to that of the rat (5). The coding sequences of both genes are interrupted at identical nucleotide positions. The introns have similar lengths except that, in the human gene, introns 12 and 16 are longer by 1.2 and 0.7 kb, respectively.

Based on alignment of the amino acid sequences of GK and the N- and C-terminal halves of rat HKII , it was postulated that the HKII gene evolved from an ancestral gene similar to glucokinase by gene duplication and tandem fusion (4, 20). The fusion point in the rat HKII was suggested to be in exon 10 which is composed of 142 bp of the tenth exon of one ancestral gene and 163 bp of the second exon from another copy of this gene (5, 20). The structural organization of the human HKII gene supports this hypothesis. The length of exons 3-9 in 5' half of the human HKII gene is identical to that of the corresponding exons 11-17 in 3' half. However, introns 2, 3, and 7 of the 5' half are significantly longer and intron 4 is significantly shorter than the corresponding introns 10, 11, 15 and 12 in the 3' half (Fig. 1). Furthermore, the organization of the human GK gene (21,22) is similar to that of the 5' half and 3' halves of the human HKII gene.

The genomic library screen for HKII clones yielded a number of clones that contained the entire coding sequence that was uninterrupted by intronic sequences. In addition, this sequence differed from that of the functional HKII cDNA sequence at several positions, including a premature stop codon at

position 120. It is very likely that this unusual sequence represents a nonfunctional processed pseudogene that originated probably by the insertion of a reverse transcript of HKII mRNA into the genome. A pseudogene of HKII was not detected in rat genomic libraries (4,5, 20).

Mutations in the gene encoding HKII, an enzyme that plays a central role in glucose metabolism, could contribute to the development of insulin resistance and NIDDM. Determination of the structural organization of the HKII gene will allow rapid screening of patients with these defects for mutations in this gene.

REFERENCES

1. Wilson, JE. (1984) in *Regulation of Carbohydrate Metabolism.* (Beitner,R. Ed.) Vol I, pp. 45-85, CRC Press Boca Raton, FL.
2. Katzen, H.M., and Schimke, R.T. (1965) *Proc. Natl. Acad. Sci. USA* 54,1218-1225.
3. Katzen, H.M., Soderman, D.D., and Cirillo, V.J. (1968) *Ann. NY. Acad. Sci. USA* 151, 351-358.
4. Thelen, A.P., and Wilson, J.E. (1991) *Arch. Biochem. Biophys.* 286, 645-51.
5. Printz, R.L., Koch, S., Potter, L.R., O'Doherty, R.M., Tiesinga, J.J., Moritz, S. and Granner, D.K. (1993) *J Biol Chem* 268, 5209-5219.
6. Deeb, S.S., Malkki, M., and Laakso, M. (1993) *Biochem and Biophys Res Comm.* 197, 68-74.
7. Schwab, D.A., and Wilson, J.E. (1989) *Proc. Natl. Acad. Sci. USA* 86, 2563-67.
8. Schwab, D.A., and Wilson, J.E. (1991) *Arch. Biochem. Biophys.* 285, 365-370.
9. Nishi, S., Seino, S., and Bell, G.I. (1988) *Biochem. Biophys. Res. Comm.* 157,937-943.
10. Easterby, J.S. and O'Brien, M.J. (1973) *Eur. J. Biochem.* 38: 201-211.
11. Ureta, T. (1982) *Comp. Biochem. Physiolog.* B 71, 549-555.
12. Lehto, M., Xiang, K., Stoffel, M., Espinosa, R., Groop, L.C., LeBeau, M.M., and Bell, G.I. (1993) *Diabetologia* 36, 1299-1302.
13. Postic, C., Leterque, A., Rencurel, F., Printz, R.L., Forest, C., Granner, D.K., and Gitrard, J. (1993) *Diabetes* 42, 922-929.
14. Granner, D.K., and O'Brien, R.M. (1992) *Diabetes Care* 15, 369-395.
15. Rothman, D.L., Shulman, R.G., and Shulman, G.I. (1992) *J. Clin. Invest.* 89, 1069-1075.
16. DelPrato, S., Bonadonna, R.C., Bonora, E., Gulli, G., Solina, A., Shank, M., and DeFronzo, R.A. (1993) *J.Clin. Invest* 91, 484-494.
17. Bell, G.I., Froguel, P., Nishi, S., Pilkis, S.J., Stoffel, M., Takeda, J., Vionnet, N., and Yasuda, K. (1993) *Trends Endocrinol. Metab.* 4, 86-90.
18. Kretz, K.A., Carson, G.S., and O'Brien, J.S. (1989) *Nucl. Acid. Res.* 17, 5864.
19. Sambrook, J.,Fritsch, E.F. and Maniatis, T. (1989) *Molecular cloning. A laboratory manual.* 2nd edition, pp. 2.63-2.81, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
20. Kogure, K., Shinohara, Y. , and Terada, H. (1993) *J. Biol. Chem.* 268, 8422-8424.
21. Stoffel, M., Froguel, P.H., Takeda, J., Zouali, H., Vionnet, N., Nishi, S., Weber, I.T., Harrison, R.W., Pilkis, S.J., Lesage, S., Vaxillaire, M., Velho, G., Sun, F., Iris, F., Passa, P.H., Cohen, D., and Bell, G.I. (1992) *Proc. Natl. Acad. Sci USA* 89: 7698-7702.
22. Tanizawa, Y., Koranyi, L.I., Welling, C.M., and Permutt, MA. (1991) *Proc. Natl. Acad. Sci. USA* 88, 7294-7297.